Semiparametric Spatiotemporal Hazard Mapping Under Local Stationarity

Jialuo Liu*, Tingjin Chu†, Jun Zhu‡ and Haonan Wang*

Colorado State University, University of Melbourne† and University of Wisconsin-Madison‡*

Example: Hazard Mapping

- The dataset consists of the measurements of noise intensity collected by 17 static sensors and 2 roving sensors.
- Measurements are taken between 10:29:00 am and 11:24:00 am when all sensors are operating.

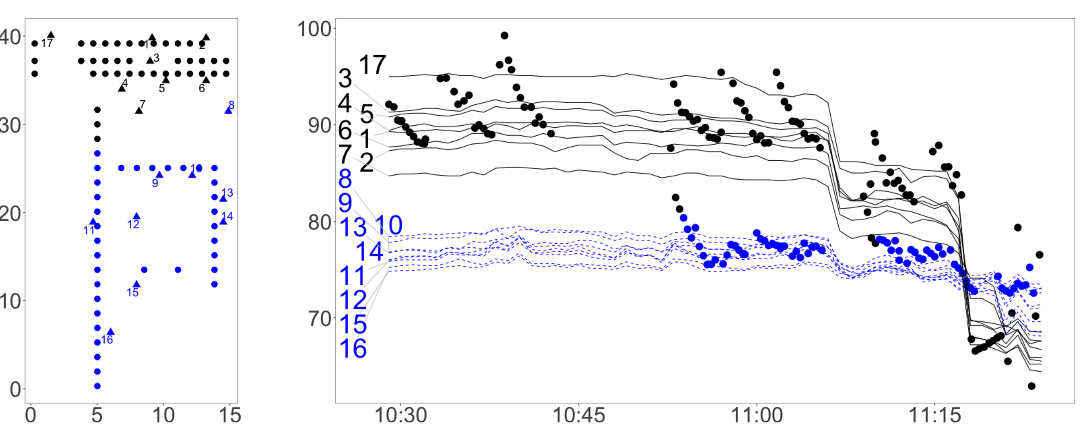


Figure 1: Locations and time series plot

- **Irregularity and Sparsity**: in space and in time
- **Nonlinearity**: nonlinear time trend
- **Nonstationarity**: inhomogeneous error variances

Spatiotemporal Sampling Design

We introduce a (L_n, T_n)-rate *spatiotemporal distance expanding asymptotics for fixed spatiotemporal domain* (STDE) asymptotics. Let $\{L_n\}, \{T_n\} \to \infty$ be two sequences of positive numbers.

- $\max_{1\le j\le N_n} \delta_{j,n} \le c_1/L_n$,
- $\max_{1\le j\le N_n} \zeta_{j,n} \le c_2/T_n$,
- $\min_{1\le j\le N_n} L_n\delta_{j,n} + T_n\zeta_{j,n} \ge c_3$,

where

- $\delta_{j,n} = \min\{\|\boldsymbol{s}_i - \boldsymbol{s}_j\| : 1 \le i \le N_n, \boldsymbol{s}_i \ne \boldsymbol{s}_j\}$
- $\zeta_{j,n} = \min\{|t_i - t_j| : 1 \le i \le N_n, t_i \ne t_j\}$

Locally Stationary Process

Consider a spatiotemporal random process $\{Y(\boldsymbol{s},t) : \boldsymbol{s} \in \mathcal{R} \subset \mathbb{R}^d, t \in \mathcal{T} \subset \mathbb{R}\}$. Denote $\mathrm{Cov}(Y(\boldsymbol{s},t), Y(\boldsymbol{s}',t')) = \gamma((\boldsymbol{s},t),(\boldsymbol{s}',t'))$. Denote the stage of the asymptotics as n.

- There exists a function $g(\boldsymbol{u}_1, u_2, \boldsymbol{s}, t)$ such that
$$\lim_{n\to\infty} \gamma_n((\boldsymbol{s},t),(\boldsymbol{s}+\boldsymbol{u}_1/L_n, t+u_2/T_n)) = g(\boldsymbol{u}_1, u_2, \boldsymbol{s}, t),$$
where $\|\boldsymbol{u}_1\| \le \tau_1, |u_2| \le \tau_2$ for any given $\tau_1 > 0, \tau_2 > 0$.
- Let $g(\boldsymbol{s},t) = g(\boldsymbol{0}, 0, \boldsymbol{s}, t)$, then $\gamma_n((\boldsymbol{s},t),(\boldsymbol{s},t)) = g(\boldsymbol{s},t) + \mathcal{O}(\rho_n)$ uniformly in $(\boldsymbol{s},t)$, where $\{\rho_n\} \to 0$ as $n \to \infty$.

Model

Consider a general Gaussian spatiotemporal model,

$$y(\boldsymbol{s},t) = \boldsymbol{x}(\boldsymbol{s},t)^\top\boldsymbol{\beta} + f(t) + \varepsilon(\boldsymbol{s},t), \quad \boldsymbol{s} \in [0,1]^d,\ t \in [0,1], \qquad (1)$$

where

- $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$.
- $f(t)$ is an unknown temporal trend function on $[0,1]$.
- $\varepsilon(\boldsymbol{s},t)$ is the spatiotemporal gaussian error process.

Generalized spatiotemporal Matérn covariance functions

We consider a class of nonseparable nonstationary S-T covariance functions

$$\gamma_n((\boldsymbol{s},t),(\boldsymbol{s}',t');\boldsymbol{\theta}) = \begin{cases} \frac{D(\boldsymbol{s},t)D(\boldsymbol{s}',t')\sigma^2(1-c)\theta_3^{d/2}2^{1-\nu}}{(\theta_1^2u_2^2+1)^\nu(\theta_1^2u_2^2+\theta_3)^{d/2}\Gamma(\nu)} m(\boldsymbol{u}_1,u_2)^\nu K_\nu\{m(\boldsymbol{u}_1,u_2)\}, & \|\boldsymbol{u}_1\|_2 > 0, \\ \frac{D(\boldsymbol{s},t)D(\boldsymbol{s}',t')\sigma^2(1-c)\theta_3^{d/2}}{(\theta_1^2u_2^2+1)^\nu(\theta_1^2u_2^2+\theta_3)^{d/2}}, & \|\boldsymbol{u}_1\|_2 = 0, |u_2| > 0, \\ D(\boldsymbol{s},t)^2\sigma^2, & \|\boldsymbol{u}_1\|_2 = 0, |u_2| = 0, \end{cases} \qquad (2)$$

where

- $\boldsymbol{u}_1 = L_n(\boldsymbol{s}-\boldsymbol{s}')$, $u_2 = T_n(t-t')$
- $m(\boldsymbol{u}_1,u_2) = \theta_2\left(\frac{\theta_1^2u_2^2+1}{\theta_1^2u_2^2+\theta_3}\right)^{1/2}\|\boldsymbol{u}_1\|$
- $K_\nu(\cdot)$ is the modified Bessel function of the second kind of order ν
- $D(\boldsymbol{s},t)$ is some known function

The above covariance function is generally nonseparable and nonstationary.

Profile Likelihood Estimation

We consider profile likelihood estimation method:

1. Let $\boldsymbol{y}^* = \boldsymbol{y} - \boldsymbol{X\beta}$. Estimate $\boldsymbol{f}$ using local polynomial regression, i.e., $\widetilde{\boldsymbol{f}} = \boldsymbol{S}\boldsymbol{y}^* = \boldsymbol{S}(\boldsymbol{y}-\boldsymbol{X\beta})$, where
$$\boldsymbol{S} = \begin{pmatrix} (1,0)(\boldsymbol{D}_{t_1}^\top\boldsymbol{K}_{t_1}\boldsymbol{D}_{t_1})^{-1}\boldsymbol{D}_{t_1}^\top\boldsymbol{K}_{t_1} \\ \vdots \\ (1,0)(\boldsymbol{D}_{t_N}^\top\boldsymbol{K}_{t_N}\boldsymbol{D}_{t_N})^{-1}\boldsymbol{D}_{t_N}^\top\boldsymbol{K}_{t_N} \end{pmatrix},$$
2. Plugging $\widetilde{\boldsymbol{f}}$ into (1), $(\boldsymbol{I}-\boldsymbol{S})\boldsymbol{y} \approx (\boldsymbol{I}-\boldsymbol{S})\boldsymbol{X\beta} + \boldsymbol{\varepsilon}$,
3. The estimates $(\widehat{\boldsymbol{\beta}}^\top, \widehat{\boldsymbol{\theta}}^\top)^\top$ minimize the following (negative) profile log-likelihood criterion $\ell(\boldsymbol{\beta},\boldsymbol{\theta})$
$$\frac{1}{2}\log|\boldsymbol{\Gamma}(\boldsymbol{\theta})| + \frac{1}{2}(\boldsymbol{y}-\boldsymbol{X\beta})^\top(\boldsymbol{I}-\boldsymbol{S})^\top\boldsymbol{\Gamma}(\boldsymbol{\theta})^{-1}(\boldsymbol{I}-\boldsymbol{S})(\boldsymbol{y}-\boldsymbol{X\beta}),$$
4. The estimates of $\boldsymbol{f}$ can be obtained as $\widehat{\boldsymbol{f}} = \boldsymbol{S}(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}})$.

Theoretical Properties

Theorem 1 *There exists, with probability tending to one, a local minimizer ${}^n\widehat{\boldsymbol{\eta}} = ({}^n\widehat{\boldsymbol{\beta}}^\top, {}^n\widehat{\boldsymbol{\theta}}^\top)^\top$ of $\ell(\boldsymbol{\eta})$ such that $\|{}^n\widehat{\boldsymbol{\eta}} - \boldsymbol{\eta}_0\| = O_p(N_n^{-1/2})$. Moreover, the local minimizer ${}^n\widehat{\boldsymbol{\eta}}$ is asymptotic normal,*

$$N_n^{1/2}({}^n\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0) \xrightarrow{D} N(\boldsymbol{0}, \boldsymbol{\Pi}^{-1}),$$
$$N_n^{1/2}({}^n\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}_0) \xrightarrow{D} N(\boldsymbol{0}, \boldsymbol{\mathcal{I}}_0(\boldsymbol{\theta}_0)^{-1}).$$

Theorem 2 *If $f^{(3)}(t)$ is bounded, under some regularity conditions, then for $t \in (0,1)$, we have*

$$(N_nh)^{1/2}\left\{\widehat{\boldsymbol{F}}(t) - \boldsymbol{F}(t) - \frac{1}{2}h^2\begin{pmatrix} \mu_2 f''(t) \\ 0 \end{pmatrix} + o(h^2)\right\} \xrightarrow{D} N\left(0, \begin{pmatrix} 1 & 0 \\ 0 & \mu_2^{-1} \end{pmatrix}\Delta_t\begin{pmatrix} 1 & 0 \\ 0 & \mu_2^{-1} \end{pmatrix}\right).$$

where $\widehat{\boldsymbol{F}}(t) = \boldsymbol{\omega}(t)(\boldsymbol{y}-\boldsymbol{X}\widehat{\boldsymbol{\beta}})$ is the estimate of $\boldsymbol{F}(t) = (f(t), hf'(t))^\top$ and $\boldsymbol{\omega}(t) = (\boldsymbol{D}_t^\top\boldsymbol{K}_t\boldsymbol{D}_t)^{-1}\boldsymbol{D}_t^\top\boldsymbol{K}_t$.

Simulation

Spatiotemporal Sampling Design:

- Consider N_s locations $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_{N_s}$ in $[0,1]^2$.
- Time points are sampled from $\{t_1, \ldots, t_{N_t}\}$, where $t_i = (i-0.5)/N_t$ and $N_t = 1000$.

Mean Structure:

- $X \sim \text{mvn}\left(\boldsymbol{0}, \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}\right)$ and $\boldsymbol{\beta} = (4,3,2,1)^\top$.
- $f(t) = 2(1-\cos(2\pi t))$.

Error Process:

- consider a special case of (2) when $\nu = 1/2$ and $\theta_3 = 1$, simplified as
$$\mathrm{Cov}(\epsilon_i,\epsilon_j) = \begin{cases} \frac{\tau^2(1-c)}{(a^2|(t_{i,N}-t_{j,N})|^2+1)^{3/2}}\exp\{-b\|\boldsymbol{s}_{i,N}-\boldsymbol{s}_{j,N}\|\}, & i \ne j \\ \tau^2, & i = j \end{cases}$$
  - $\tau^2 = D(\boldsymbol{s_i},t_i)D(\boldsymbol{s_j},t_j)\sigma^2$
  - $\boldsymbol{s}_{i,N} = L_N\boldsymbol{s}_i$ and $t_{i,N} = T_Nt_i$
  - $D(\boldsymbol{s_i},t_i) = dt_i + 1$
  - $\sigma^2 = 0.2, c = 0.2, a = 1, b = 1, d = 2$

We consider three sample sizes $N_n = 806, 1644, 2449$ by letting $N_s = 20, 40, 60$ and the following results are summarized from 400 iterations.

N_s	$N_s = 20$		$N_s = 40$		$N_s = 60$	
h	0.032		0.029		0.027	
Regression parameters						
β_1	3.996(0.032)	0.032	4.001(0.023)	0.023	4.001(0.019)	0.019
β_2	3.003(0.037)	0.034	3.002(0.022)	0.023	2.999(0.018)	0.018
β_3	2.003(0.032)	0.033	1.997(0.023)	0.024	2.000(0.019)	0.018
β_4	1.000(0.033)	0.032	0.999(0.022)	0.023	0.999(0.018)	0.018
Covariance parameters						
σ^2	0.210(0.029)	0.028	0.205(0.020)	0.020	0.203(0.016)	0.016
c	0.206(0.156)	0.138	0.213(0.093)	0.080	0.196(0.076)	0.068
a	0.947(0.296)	0.226	0.939(0.179)	0.135	0.987(0.143)	0.117
b	0.995(0.298)	0.298	0.970(0.146)	0.159	0.981(0.110)	0.124
d	1.967(0.279)	0.274	1.989(0.204)	0.197	1.994(0.156)	0.159
MSPE	0.824		0.718		0.841	

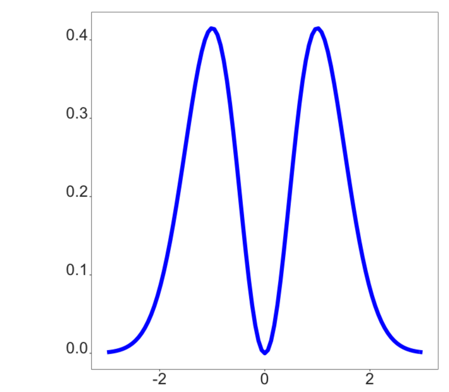
Table 1: Parameter estimates.

Bandwidth Selection

Theorem 3 *For the spatiotemporal partial linear model (1) with a locally stationary covariance function, we have*

$$E\{CV(h)\} = \frac{1}{N_n}\sum_{i=1}^{N_n} E\{f(t_i) - \hat{f}^{(-i)}(t_i)\}^2 + \overline{\sigma^2} + o\left(\frac{1}{N_nh}\right) - \frac{K(0)}{N_nhq(t) - K(0)}\left(\frac{2}{N_n}\sum_{i=1}^{N_n}\sum_{\substack{j\ne i \\ t_j = t_i}}\gamma(i,j)\right),$$

- $K(0) = 0$ will remove the effect of correlated errors.
- Bimodal kernel: $K_2(u) = \frac{2}{\sqrt{\pi}}u^2\exp(-u^2)$.



- Bandwidth selection criterion: cross-validation (CV)

Real Data Analysis

For the noise intensity data, we have the fitted mean structure

$$\widehat{y} = -0.51s_1 + 0.45s_2 + \widehat{f}(t)$$

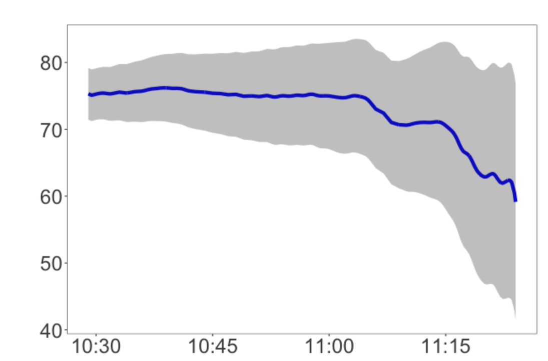


Figure 2: $\widehat{f}(t)$

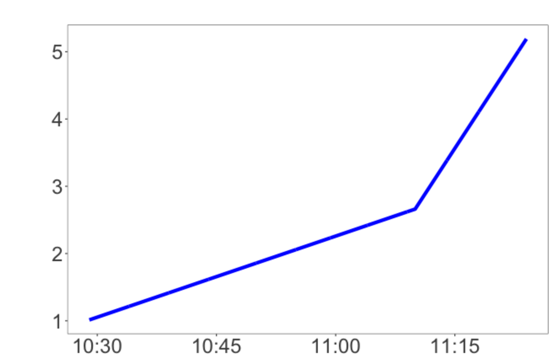


Figure 3: $D(\boldsymbol{s},t)$

Here, $D(\boldsymbol{s},t) = 1 + 2.22t + 7.77(t-\tau)_+$, and τ is chosen as 11:10:00.

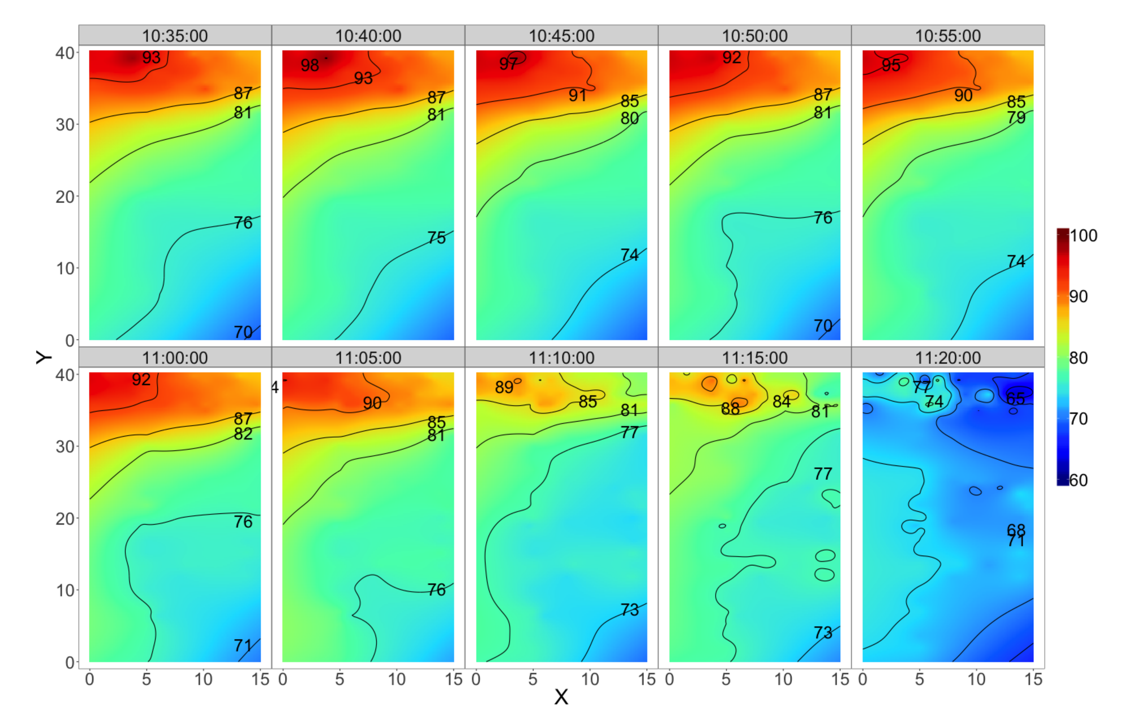


Figure 4: Dynamic hazard maps

- The dynamic hazard maps suggest a possible noise source on the upper-left corner.
- A change of the overall noise intensity at 11:10:00.
- A horizontal separation around $y = 30$ before 11:10:00.